Relation Extraction

Luke Zettlemoyer

CSE 517

Winter 2013

[with slides adapted from many people, including Bill MacCartney, Dan Jurafsky, Rion Snow, Jim Martin, Chris Manning, William Cohen, and others]

Goal: “machine reading”

- Acquire structured knowledge from unstructured text

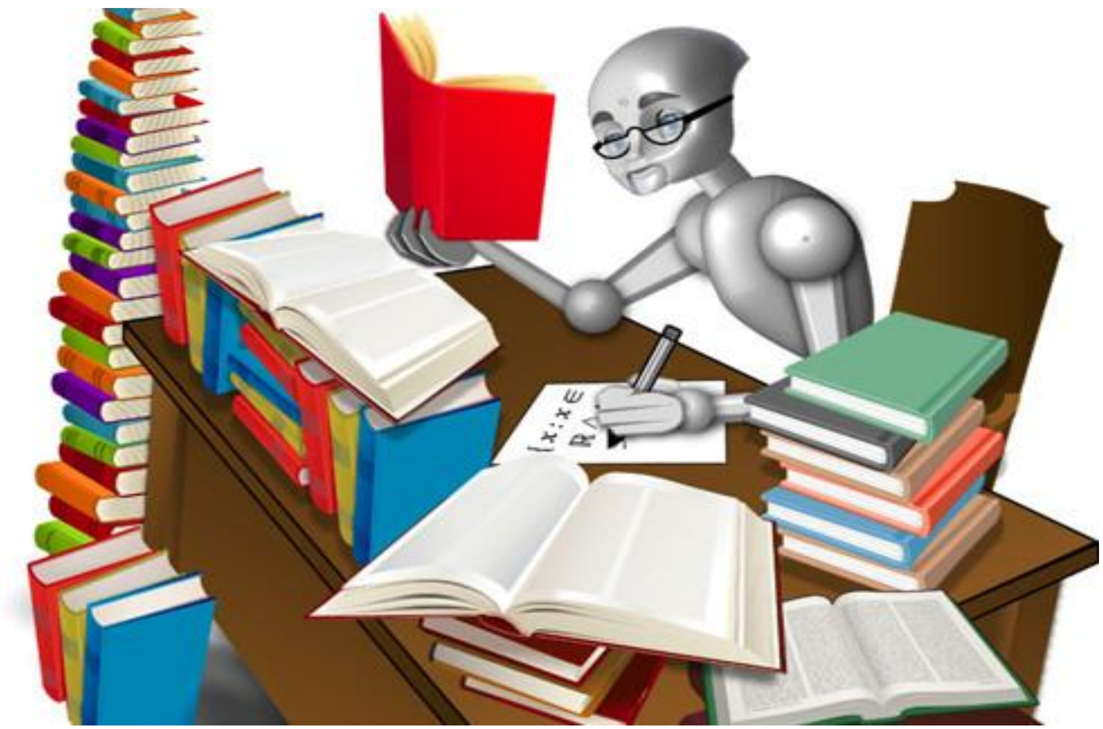


illustration from DARPA

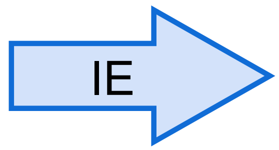
Information extraction

- IE = extracting information from text
- Sometimes called *text analytics* commercially
- Extract entities
  - People, organizations, locations, times, dates, prices, ...
  - Or sometimes: genes, proteins, diseases, medicines, ...
- Extract the relations between entities
  - Located in, employed by, part of, married to, ...
- Figure out the larger events that are taking place

Machine-readable summaries


Involvement of Tumor Necrosis Factor Receptor-associated Protein 1 (TRAP1) in Apoptosis Induced by β-Hydroxyisovalerylshikonin*

textual abstract:
summary for human



Subject	Relation	Object
p53	**is_a**	protein
Bax	**is_a**	protein
p53	has_function	apoptosis
Bax	has_function	induction
apoptosis	involved_in	cell_death
Bax	is_in	mitochondrial outer membrane
Bax	is_in	cytoplasm
apoptosis	related_to	caspase activation
...	...	...

structured knowledge extraction:
summary for machine

More applications of IE

- Building & extending knowledge bases and ontologies
- Scholarly literature databases: Google Scholar, CiteSeerX
- People directories: Rapleaf, Spoke, Naymz
- Shopping engines & product search
- Bioinformatics: clinical outcomes, gene interactions, …
- Patent analysis
- Stock analysis: deals, acquisitions, earnings, hirings & firings
- SEC filings
- Intelligence analysis for business & government

Named Entity Recognition (NER)

The task:

1. find names in text
2. classify them by type, usually {ORG, PER, LOC, MISC}

```
The [European Commission ORG] said on Thursday it
disagreed with [German MISC] advice.
Only [France LOC] and [Britain LOC] backed
[Fischler PER] 's proposal .

"What we have to be extremely careful of is how
other countries are going to take [Germany LOC]
's lead", [Welsh National Farmers ' Union ORG]
( [NFU ORG] ) chairman [John Lloyd Jones PER]
said on [BBC ORG] radio .
```

Named Entity Recognition (NER)

- It's a tagging task, similar to part-of speech (POS) tagging
- So, systems use sequence classifiers: HMMs, MEMMs, CRFs
- Features usually include words, POS tags, word shapes, orthographic features, gazetteers, etc.
- Accuracies of >90% are typical — but depends on genre!
- NER is commonly thought of as a ”solved problem”
- A building block technology for relation extraction
- E.g., http://nlp.stanford.edu/software/CRF-NER.shtml

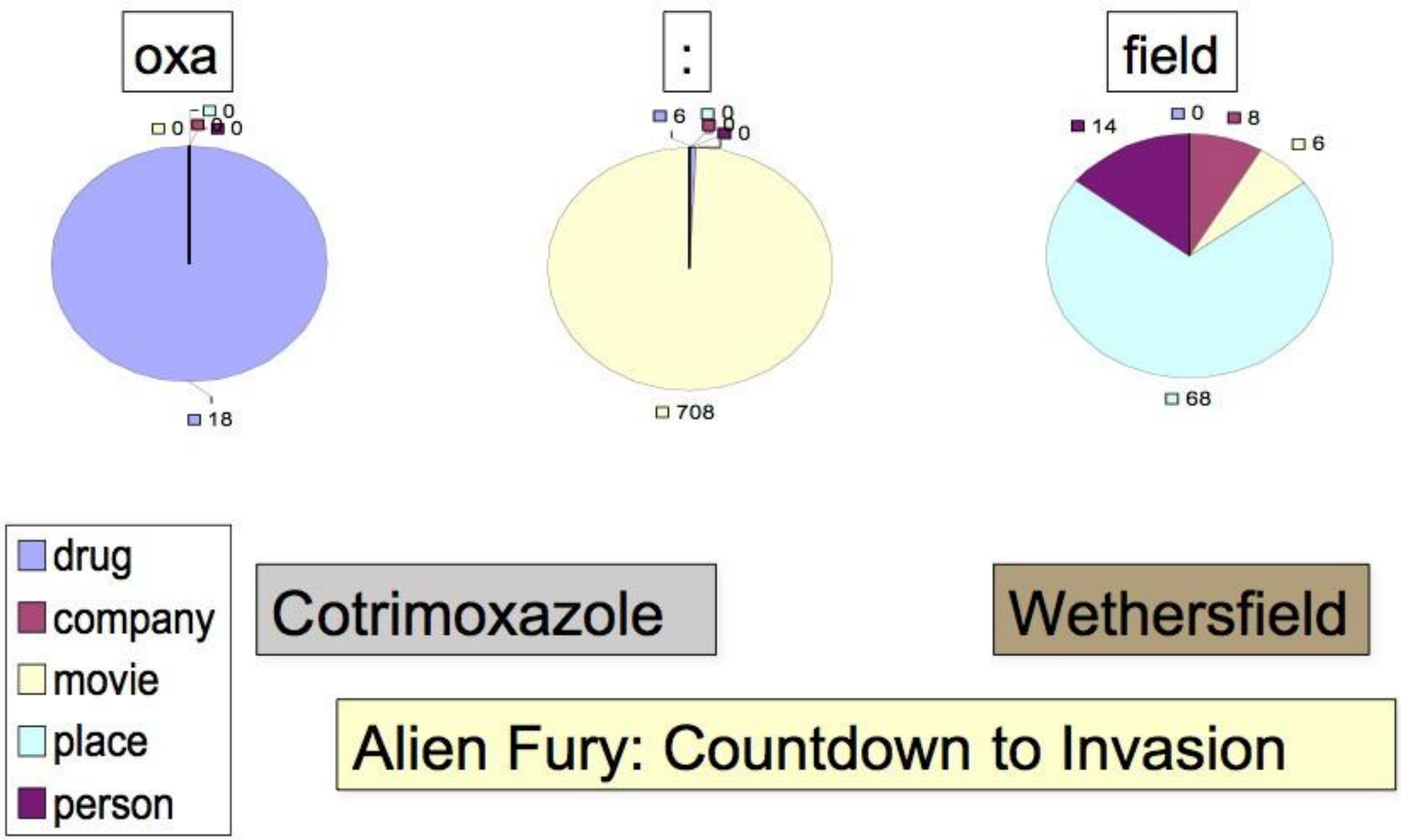
Orthographic features for NER

oxa : field

drug
company
movie
place
person

slide adapted from Chris Manning

Orthographic features for NER



Relation extraction example

CHICAGO (AP) — Citing high fuel prices, United Airlines said Friday it has increased fares by $6 per round trip on flights to some cities also served by lower-cost carriers. American Airlines, a unit of AMR, immediately matched the move, spokesman Tim Wagner said. United, a unit of UAL, said the increase took effect Thursday night and applies to most routes where it competes against discount carriers, such as Chicago to Dallas and Atlanta and Denver to San Francisco, Los Angeles and New York.

Question: What relations should we extract?

example from Jim Martin

Relation extraction example

CHICAGO (AP) — Citing high fuel prices, United Airlines said Friday it has increased fares by $6 per round trip on flights to some cities also served by lower-cost carriers. American Airlines, a unit of AMR, immediately matched the move, spokesman Tim Wagner said. United, a unit of UAL, said the increase took effect Thursday night and applies to most routes where it competes against discount carriers, such as Chicago to Dallas and Atlanta and Denver to San Francisco, Los Angeles and New York.

Subject	Relation	Object
American Airlines	subsidiary	AMR
Tim Wagner	employee	American Airlines
United Airlines	subsidiary	UAL

example from Jim Martin

Relation types

For generic news texts ...

Relations		Examples	Types
Affiliations			
	Personal	*married to, mother of*	PER $\rightarrow$ PER
	Organizational	*spokesman for, president of*	PER $\rightarrow$ ORG
	Artifactual	*owns, invented, produces*	(PER $\mid$ ORG) $\rightarrow$ ART
Geospatial			
	Proximity	*near, on outskirts*	LOC $\rightarrow$ LOC
	Directional	*southeast of*	LOC $\rightarrow$ LOC
Part-Of			
	Organizational	*a unit of, parent of*	ORG $\rightarrow$ ORG
	Political	*annexed, acquired*	GPE $\rightarrow$ GPE

slide adapted from Jim Martin

Relation types from ACE 2003

ROLE: relates a person to an organization or a geopolitical entity
subtypes: member, owner, affiliate, client, citizen

PART: generalized containment
subtypes: subsidiary, physical part-of, set membership

AT: permanent and transient locations
subtypes: located, based-in, residence

SOCIAL: social relations among persons
subtypes: parent, sibling, spouse, grandparent, associate

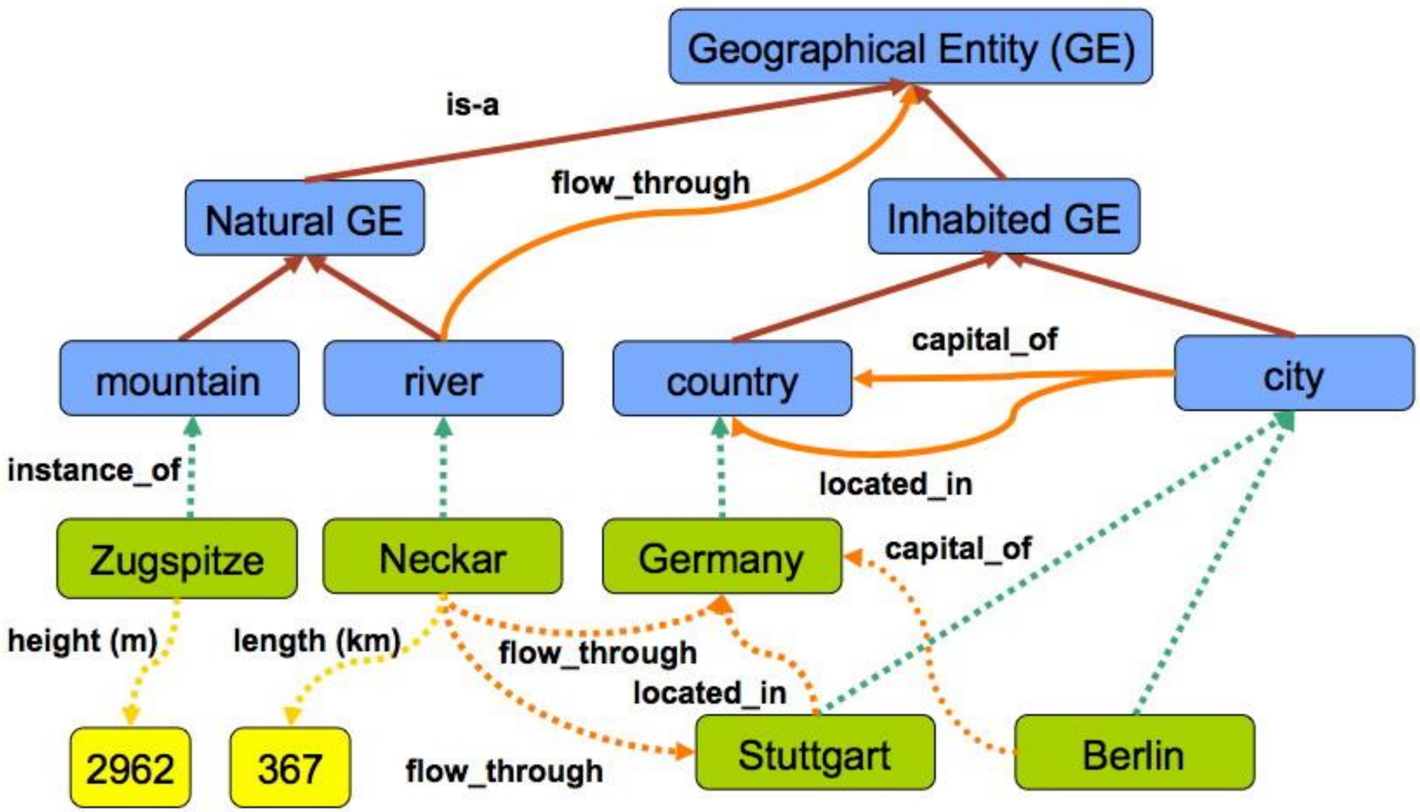
slide adapted from Doug Appelt

Relation types: Freebase

23 Million Entities, thousands of relations

Relation name	Size	Example
/people/person/nationality	281,107	John Dugard, South Africa
/location/location/contains	253,223	Belgium, Nijlen
/people/person/profession	208,888	Dusa McDuff, Mathematician
/people/person/place_of_birth	105,799	Edwin Hubble, Marshfield
/dining/restaurant/cuisine	86,213	MacAyo's Mexican Kitchen, Mexican
/business/business_chain/location	66,529	Apple Inc., Apple Inc., South Park, NC
/biology/organism_classification_rank	42,806	Scorpaeniformes, Order
/film/film/genre	40,658	Where the Sidewalk Ends, Film noir
/film/film/language	31,103	Enter the Phoenix, Cantonese
/biology/organism_higher_classification	30,052	Calopteryx, Calopterygidae
/film/film/country	27,217	Turtle Diary, United States
/film/writer/film	23,856	Irving Shulman, Rebel Without a Cause
/film/director/film	23,539	Michael Mann, Collateral
/film/producer/film	22,079	Diane Eskenazi, Aladdin
/people/deceased_person/place_of_death	18,814	John W. Kern, Asheville
/music/artist/origin	18,619	The Octopus Project, Austin
/people/person/religion	17,582	Joseph Chartrand, Catholicism
/book/author/works_written	17,278	Paul Auster, Travels in the Scriptorium
/soccer/football_position/players	17,244	Midfielder, Chen Tao
/people/deceased_person/cause_of_death	16,709	Richard Daintree, Tuberculosis
/book/book/genre	16,431	Pony Soldiers, Science fiction
/film/film/music	14,070	Stavisky, Stephen Sondheim
/business/company/industry	13,805	ATS Medical, Health care

Relation types: geographical



slide adapted from Paul Buitelaar

More relations: disease outbreaks

May 19 1995, Atlanta -- The Centers for Disease Control and Prevention, which is in the front line of the world's response to the deadly **Ebola** epidemic in **Zaire**, is finding itself hard pressed to cope with the crisis…

Information Extraction System (e.g., NYU's Proteus)

Disease Outbreaks in ***The New York Times***

Date	***Disease Name***	***Location***
Jan. 1995	Malaria	Ethiopia
July 1995	Mad Cow Disease	U.K.
Feb. 1995	Pneumonia	U.S.

slide adapted from Eugene Agichtein

More relations: protein interactions

„We show that CBF-A and CBF-C interact with each other to form a CBF-A-CBF-C complex and that CBF-B does not interact with CBF-A or CBF-C individually but that it associates with the CBF-A-CBF-C complex."

$$\text{CBF-A} \xleftrightarrow[\text{complex}]{\text{interact}} \text{CBF-C}$$

$$\text{CBF-B} \xrightarrow{\text{associates}} \text{CBF-A-CBF-C complex}$$

slide adapted from Rosario & Hearst

Relations between word senses

- NLP applications need word meaning!
  - Question answering
  - Conversational agents
  - Summarization
- One key meaning component: word relations
  - Hyponymy: San Francisco is an instance of a city
  - Antonymy: acidic is the opposite of basic
  - Meronymy: an alternator is a part of a car

WordNet is incomplete

Ontological relations are missing for many words:

In WordNet 3.1	Not in WordNet 3.1
insulin progesterone	leptin pregnenolone
combustibility navigability	affordability reusability
HTML	XML
Google, Yahoo	Microsoft, IBM

Esp. for specific domains: restaurants, auto parts, finance

Relation extraction: 5 easy methods

1. Hand-built patterns
2. Bootstrapping methods
3. Supervised methods
4. Distant supervision
5. Unsupervised methods

Relation extraction: 5 easy methods

1. Hand-built patterns
2. Bootstrapping methods
3. Supervised methods
4. Distant supervision
5. Unsupervised methods

A hand-built extraction rule

```
;;; For <company> appoints <person> <position>

(defpattern appoint
   "np-sem(C-company)?  rn?  sa?  vg(C-appoint) np-sem(C-person) ´,´?
    to-be?  np(C-position) to-succeed?:
    company-at=1.attributes, sa=3.span, lv=4.span, person-at=5.attribute:
    position-at=8.attributes |
...
```

```
(defun when-appoint (phrase-type)
    (let ((person-at (binding ´person-at))
         (company-entity (entity-bound ´company-at))
         (person-entity (essential-entity-bound ´person-at ´C-person))
         (position-entity (entity-bound ´position-at))
         (predecessor-entity (entity-bound ´predecessor-at))
         new-event)
     (not-an-antecedent position-entity)
     ;; if no company is specified for position, use agent
...
```

NYU Proteus system (1997)

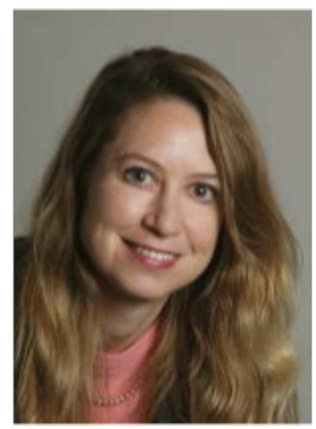
Patterns for learning hyponyms

- Intuition from Hearst (1992)

  Agar is a substance prepared from a mixture of red algae, such as Gelidium, for laboratory or industrial use.

- What does *Gelidium* mean?

- How do you know?



Patterns for learning hyponyms

- Intuition from Hearst (1992)

  Agar is a substance prepared from a mixture of red algae, such as Gelidium, for laboratory or industrial use.

- What does *Gelidium* mean?

- How do you know?

Hearst’s lexico-syntactic patterns

Y such as X ((, X)* (, and/or) X)
such Y as X…
X… or other Y
X… and other Y
Y including X…
Y, especially X…

Hearst, 1992. Automatic Acquisition of Hyponyms.

Examples of the Hearst patterns

Hearst pattern	Example occurrences
X and other Y	...temples, treasuries, and other important civic buildings.
X or other Y	bruises, wounds, broken bones or other injuries...
Y such as X	The bow lute, such as the Bambara ndang...
such Y as X	...such authors as Herrick, Goldsmith, and Shakespeare.
Y including X	...common-law countries, including Canada and England...
Y, especially X	European countries, especially France, England, and Spain...

Patterns for learning meronyms

- Berland & Charniak (1999) tried it
- Selected initial patterns by finding all sentences in a corpus containing *basement* and *building*


whole NN[-PL] 's POS part NN[-PL]	... building's basement ...
part NN[-PL] of PREP {the|a} DET mods [JJ|NN]* whole NN	... basement of a building ...
part NN in PREP {the|a} DET mods [JJ|NN]* whole NN	... basement in a building ...
parts NN-PL of PREP wholes NN-PL	... basements of buildings ...
parts NN-PL in PREP wholes NN-PL	... basements in buildings ...

- Then, for each pattern:
  1. found occurrences of the pattern
  2. filtered those ending with *-ing*, *-ness*, *-ity*
  3. applied a likelihood metric — poorly explained
- Only the first two patterns gave decent (though not great!) results

Problems with hand-built patterns

- Requires hand-building patterns for each relation!
  - hard to write; hard to maintain
  - there are zillions of them
  - domain-dependent

- Don't want to do this for all possible relations!

- Plus, we'd like better accuracy
  - Hearst: 66% accuracy on hyponym extraction
  - Berland & Charniak: 55% accuracy on meronyms

Relation extraction: 5 easy methods

1. Hand-built patterns
2. Bootstrapping methods
3. Supervised methods
4. Distant supervision
5. Unsupervised methods

Bootstrapping approaches

- If you don't have enough annotated text to train on …
- But you do have:
  - some seed instances of the relation
  - (or some patterns that work pretty well)

  - and lots & lots of unannotated text (e.g., the web)

- … can you use those seeds to do something useful?
- Bootstrapping can be considered *semi-supervised*

Bootstrapping example

- Target relation: *burial place*
- Seed tuple: [*Mark Twain*, *Elmira*]
- Grep/Google for "Mark Twain" and "Elmira"

  "Mark Twain is buried in Elmira, NY."

  → X is buried in Y

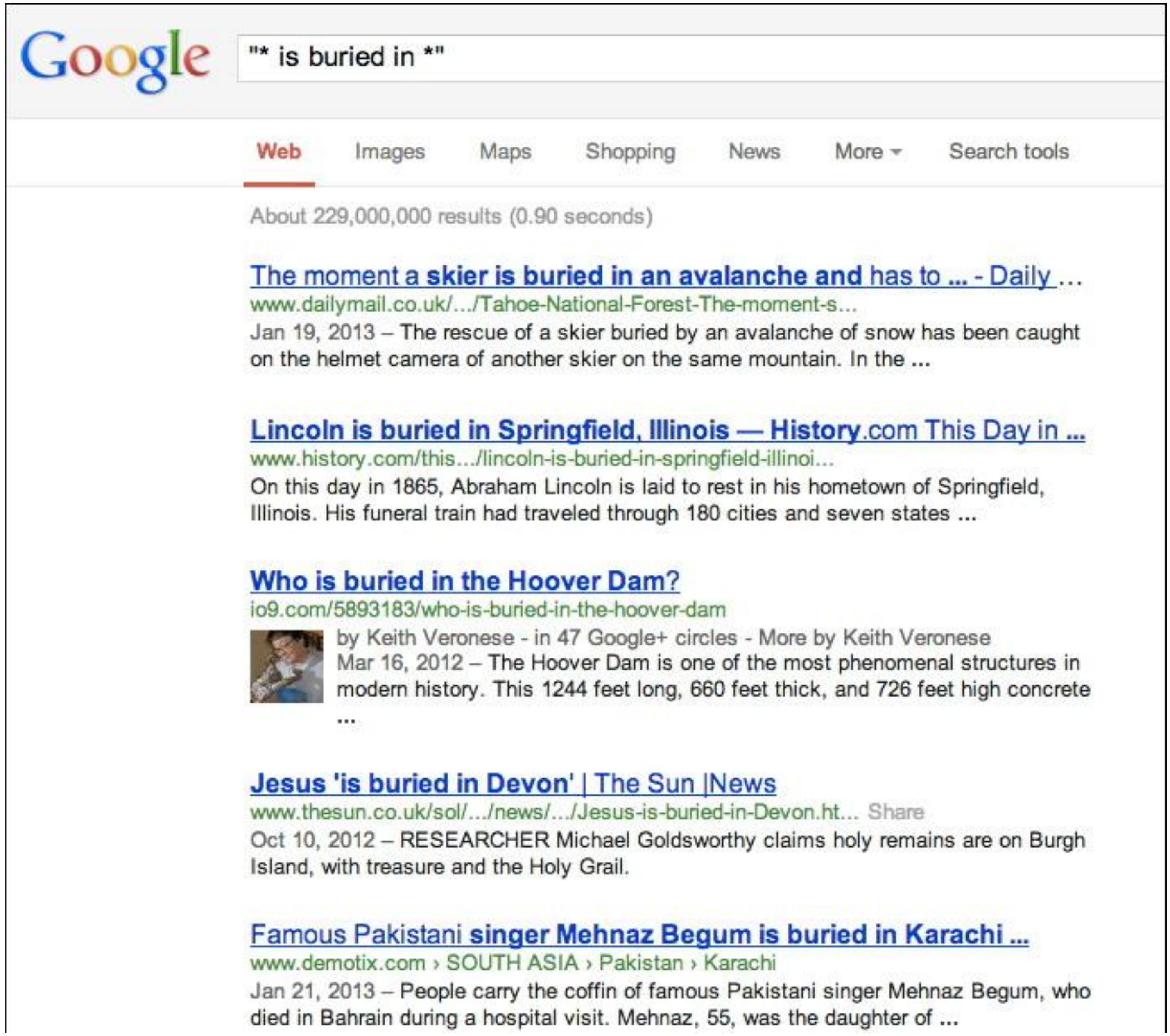
  "The grave of Mark Twain is in Elmira"

  → The grave of X is in Y

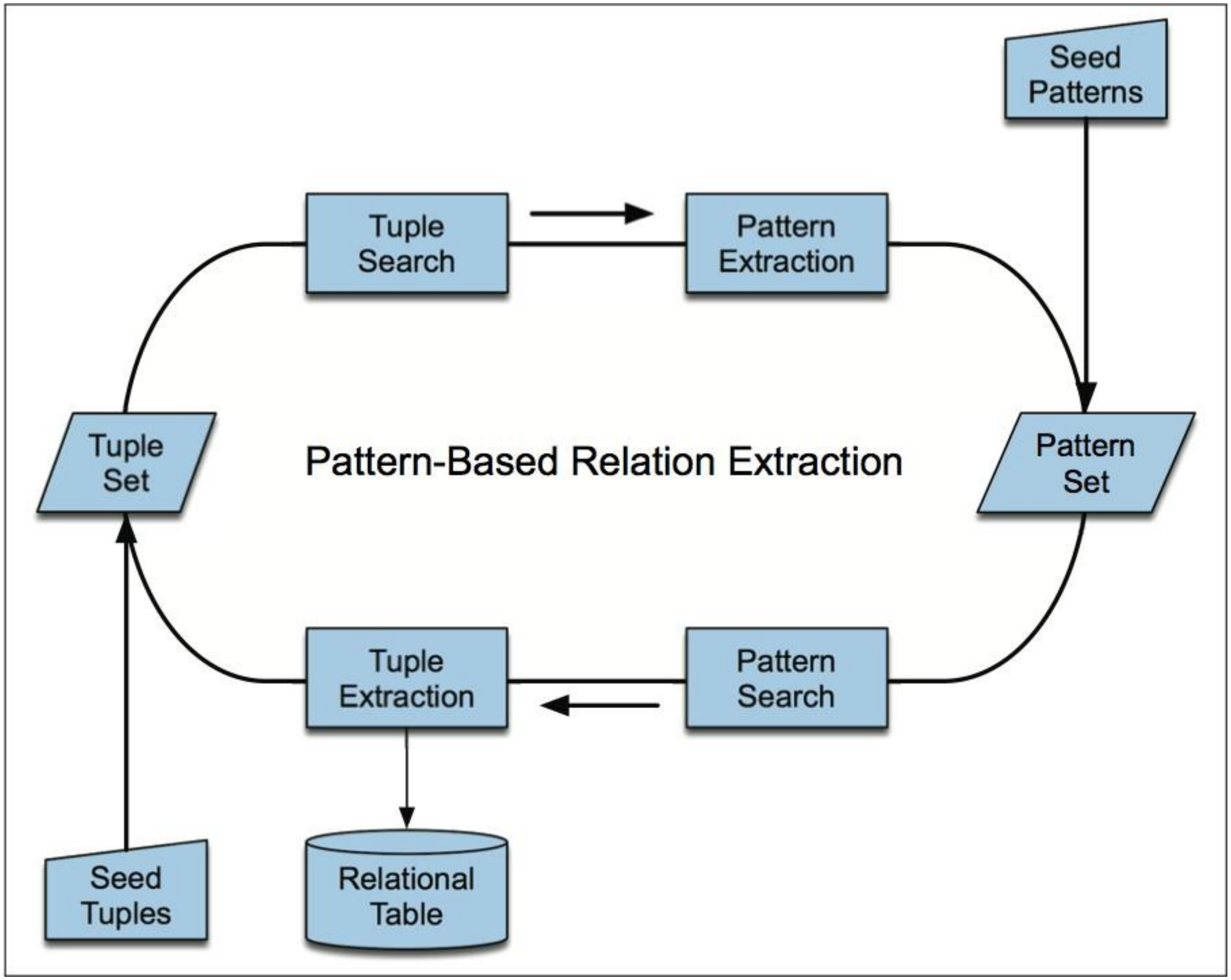
  "Elmira is Mark Twain's final resting place"

  → Y is X's final resting place
- Use those patterns to search for new tuples

Bootstrapping example



Bootstrapping relations



slide adapted from Jim Martin

DIPRE (Brin 1998)

Extract (author, book) pairs

Start with these 5 seeds:

Author	Book
Isaac Asimov	The Robots of Dawn
David Brin	Startide Rising
James Gleick	Chaos: Making a New Science
Charles Dickens	Great Expectations
William Shakespeare	The Comedy of Errors

Learn these patterns:

URL Prefix	Text Pattern
`www.sff.net/locus/c.*`	<LI><B>*title*</B> by *author* (
`dns.city-net.com/~lmann/awards/hugos/1984.html`	<i>*title*</i> by *author* (
`dolphin.upenn.edu/~dcummins/texts/sf-award.htm`	*author* $\|$ *title* $\|$ (

Iterate: use patterns to get more instances & patterns…

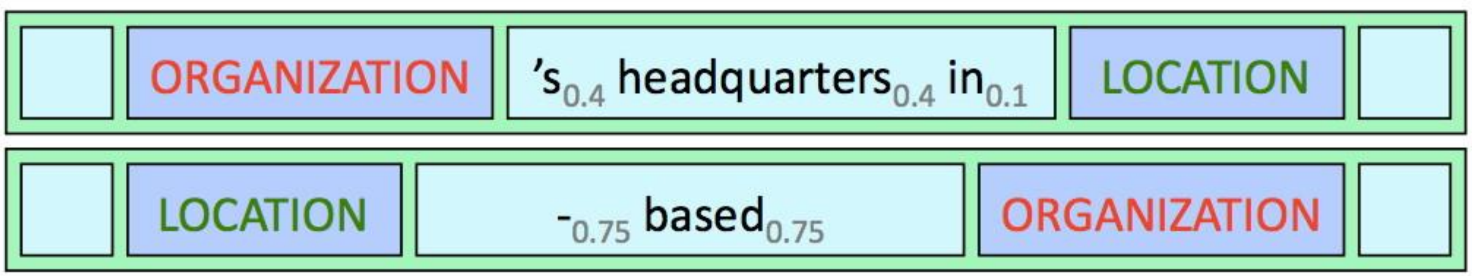
Results: after three iterations of bootstrapping loop, extracted 15,000 author-book pairs with 95% accuracy.

Snowball (Agichtein & Gravano 2000)

New ideas:

- require that X and Y be named entities
- add heuristics to score extractions, select best ones

Organization	Location of Headquarters
Microsoft	Redmond
Exxon	Irving
IBM	Armonk
Boeing	Seattle
Intel	Santa Clara



Snowball Results!

Conf	*middle*	*right*
1	<based, 0.53> <in, 0.53>	<, , 0.01>
0.69	<', 0.42> <s, 0.42> < headquarters, 0.42> <in, 0.12>	
0.61	<(, 0.93>	<), 0.12>

Table 2: Actual patterns discovered by *Snowball*. (For each pattern the *left* vector is empty, *tag1* = *ORGANIZATION*, and *tag2* = *LOCATION*.)

			Type of Error			
	Correct	*Incorrect*	*Location*	*Organization*	*Relationship*	P_{Ideal}
DIPRE	74	26	3	18	5	90%
Snowball (all tuples)	52	48	6	41	1	88%
Snowball ($\tau_t = 0.8$)	93	7	3	4	0	96%
Baseline	25	75	8	62	5	66%

5: Manually computed precision estimate, derived from a random sample of 100 tuples from each ex

Bootstrapping problems

- Requires that we have seeds for each relation
  - Sensitive to original set of seeds
- Big problem of semantic drift at each iteration
- Precision tends to be not that high
- Generally have lots of parameters to be tuned
- No probabilistic interpretation
  - Hard to know how confident to be in each result

Relation extraction: 5 easy methods

1. Hand-built patterns
2. Bootstrapping methods
3. Supervised methods
4. Distant supervision
5. Unsupervised methods

Supervised relation extraction

The supervised approach requires:

- Defining an inventory of output labels
  - Relation detection: true/false
  - Relation classification: located-in, employee-of, inventor-of, …
- Collecting labeled training data: MUC, ACE, …
- Defining a feature representation: words, entity types, …
- Choosing a classifier: Naïve Bayes, MaxEnt, SVM, …
- Evaluating the results

ACE 2008: relations

Type	Subtype
ART (artifact)	User-Owner-Inventor-Manufacturer
GEN-AFF (General affiliation)	Citizen-Resident-Religion-Ethnicity, Org-Location
METONYMY*	*None*
ORG-AFF (Org-affiliation)	Employment, Founder, Ownership, Student-Alum, Sports-Affiliation, Investor-Shareholder, Membership
PART-WHOLE (part-to-whole)	Artifact, Geographical, Subsidiary
PER-SOC* (person-social)	Business, Family, Lasting-Personal
PHYS* (physical)	Located, Near

ACE 2008: data

Source	Training epoch	Approximate size
English Resources		
Broadcast News	3/03 – 6/03	55,000 words
Broadcast Conversations	3/03 – 6/03	40,000 words
Newswire	3/03 – 6/03	50,000 words
Weblog	11/04 – 2/05	40,000 words
Usenet	11/04 – 2/05	40,000 words
Conversational Telephone Speech	11/04-12/04 (differentiated by topic vs. eval)	40,000 words
Arabic Resources		
Broadcast News	10/00 – 12/00	30,000+ words
Newswire	10/00 – 12/00	55,000+ words
Weblog	11/04 – 2/05	20,000+ words

Features

- Lightweight features — require little pre-processing
  - Bags of words & bigrams between, before, and after the entities
  - Stemmed versions of the same
  - The types of the entities
  - The distance (number of words) between the entities
- Medium-weight features — require base phrase chunking
  - Base-phrase chunk paths
  - Bags of chunk heads
- Heavyweight features — require full syntactic parsing
  - Dependency-tree paths
  - Constituent-tree paths
  - Tree distance between the entities
  - Presence of particular constructions in a constituent structure

Let’s take a close look at features used in Zhou et al. 2005

Features: words

American Airlines, a unit of AMR, immediately matched the move, spokesman Tim Wagner said.

Bag-of-words features

WM1 = {American, Airlines}, WM2 = {Tim, Wagner}

Head-word features

HM1 = Airlines, HM2 = Wagner, HM12 = Airlines+Wagner

Words in between

WBNULL = false, WBFL = NULL, WBF = a, WBL = spokesman,
WBO = {unit, of, AMR, immediately, matched, the, move}

Words before and after

BM1F = NULL, BM1L = NULL, AM2F = said, AM2L = NULL

Word features yield good precision (69%), but poor recall (24%)

Features: NE type & mention level

American Airlines, a unit of AMR, immediately matched the move, spokesman Tim Wagner said.

Named entity types (ORG, LOC, PER, etc.)
ET1 = ORG, ET2 = PER, ET12 = ORG-PER

Mention levels (NAME, NOMINAL, or PRONOUN)
ML1 = NAME, ML2 = NAME, ML12 = NAME+NAME

Named entity type features help recall a lot (+8%)
Mention level features have little impact

Features: overlap

American Airlines, a unit of AMR, immediately matched the move, spokesman Tim Wagner said.

Number of mentions and words in between

#MB = 1, #WB = 9

Does one mention include in the other?

M1>M2 = false, M1<M2 = false

Conjunctive features

ET12+M1>M2 = ORG-PER+false

ET12+M1<M2 = ORG-PER+false

HM12+M1>M2 = Airlines+Wagner+false

HM12+M1<M2 = Airlines+Wagner+false

These features hurt precision a lot (-10%), but also help recall a lot (+8%)

Features: base phrase chunking

American Airlines, a unit of AMR, immediately matched the move, spokesman Tim Wagner said.

Parse using the Stanford Parser, then apply Sabine Buchholz's chunklink.pl:

```
 0 B-NP    NNP   American     NOFUNC   Airlines     1 B-S/B-S/B-NP/B-NP
 1 I-NP    NNPS  Airlines     NP       matched      9 I-S/I-S/I-NP/I-NP
 2 O       COMMA COMMA        NOFUNC   Airlines     1 I-S/I-S/I-NP
 3 B-NP    DT    a            NOFUNC   unit         4 I-S/I-S/I-NP/B-NP/B-NP
 4 I-NP    NN    unit         NP       Airlines     1 I-S/I-S/I-NP/I-NP/I-NP
 5 B-PP    IN    of           PP       unit         4 I-S/I-S/I-NP/I-NP/B-PP
 6 B-NP    NNP   AMR          NP       of           5 I-S/I-S/I-NP/I-NP/I-PP/B-NP
 7 O       COMMA COMMA        NOFUNC   Airlines     1 I-S/I-S/I-NP
 8 B-ADVP  RB    immediately  ADVP     matched      9 I-S/I-S/B-ADVP
 9 B-VP    VBD   matched      VP/S     matched      9 I-S/I-S/B-VP
10 B-NP    DT    the          NOFUNC   move        11 I-S/I-S/I-VP/B-NP
11 I-NP    NN    move         NP       matched      9 I-S/I-S/I-VP/I-NP
12 O       COMMA COMMA        NOFUNC   matched      9 I-S
13 B-NP    NN    spokesman    NOFUNC   Wagner      15 I-S/B-NP
14 I-NP    NNP   Tim          NOFUNC   Wagner      15 I-S/I-NP
15 I-NP    NNP   Wagner       NP       matched      9 I-S/I-NP
16 B-VP    VBD   said         VP       matched      9 I-S/B-VP
17 O       .     .            NOFUNC   matched      9 I-S
```

[$_{NP}$ American Airlines], [$_{NP}$ a unit] [$_{PP}$ of] [$_{NP}$ AMR], [$_{ADVP}$ immediately] [$_{VP}$ matched] [$_{NP}$ the move], [$_{NP}$ spokesman Tim Wagner] [$_{VP}$ said].

Features: base phrase chunking

[$_{NP}$ American Airlines], [$_{NP}$ a unit] [$_{PP}$ of] [$_{NP}$ AMR], [$_{ADVP}$ immediately] [$_{VP}$ matched] [$_{NP}$ the move], [$_{NP}$ spokesman Tim Wagner] [$_{VP}$ said].

Phrase heads before and after

CPHBM1F = NULL, CPHBM1L = NULL, CPHAM2F = said, CPHAM2L = NULL

Phrase heads in between

CPHBNULL = false, CPHBFL = NULL, CPHBF = unit, CPHBL = move

CPHBO = {of, AMR, immediately, matched}

Phrase label paths

CPP = [NP, PP, NP, ADVP, VP, NP]

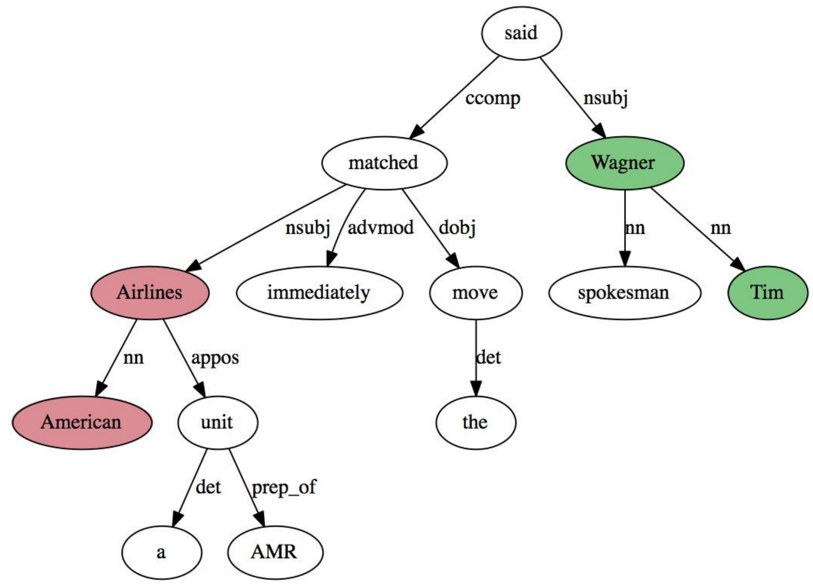
CPPH = NULL

These features increased both precision & recall by 4-6%

Features: syntactic features

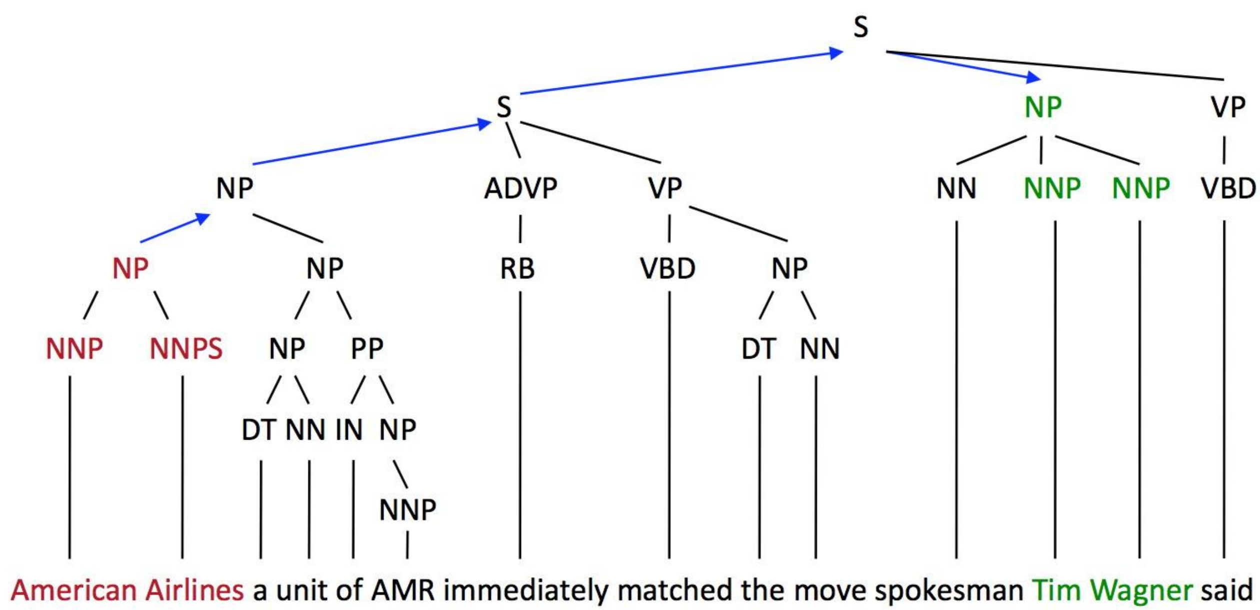
Features of mention dependencies
ET1DW1 = ORG:Airlines
H1DW1 = matched:Airlines
ET2DW2 = PER:Wagner
H2DW2 = said:Wagner

Features describing entity types and dependency tree
ET12SameNP = ORG-PER-false
ET12SamePP = ORG-PER-false
ET12SameVP = ORG-PER-false



These features had disappointingly little impact!

Features: syntactic features



Phrase label paths

PTP = [NP, S, NP]

PTPH = [NP:Airlines, S:matched, NP:Wagner]

These features had disappointingly little impact!

Relation extraction classifiers

Now use any (multiclass) classifier you like:

- SVM
- MaxEnt (aka multiclass logistic regression)
- Naïve Bayes
- etc.

[Zhou et al. 2005 used a one-vs-many SVM]

Zhou et al. 2005 results

Features	P	R	F
Words	69.2	23.7	35.3
+Entity Type	67.1	32.1	43.4
+Mention Level	67.1	33.0	44.2
+Overlap	57.4	40.9	47.8
+Chunking	61.5	46.5	53.0
+Dependency Tree	62.1	47.2	53.6
+Parse Tree	62.3	47.6	54.0
+Semantic Resources	63.1	49.5	55.5

Table 2: Contribution of different features over 43 relation subtypes in the test data

Zhou et al. 2005 results

Type	Subtype	#Testing Instances	#Correct	#Error	P	R	F
AT		**392**	**224**	**105**	**68.1**	**57.1**	**62.1**
	Based-In	85	39	10	79.6	45.9	58.2
	Located	241	132	120	52.4	54.8	53.5
	Residence	66	19	9	67.9	28.8	40.4
NEAR		**35**	**8**	**1**	**88.9**	**22.9**	**36.4**
	Relative-Location	35	8	1	88.9	22.9	36.4
PART		**164**	**106**	**39**	**73.1**	**64.6**	**68.6**
	Part-Of	136	76	32	70.4	55.9	62.3
	Subsidiary	27	14	23	37.8	51.9	43.8
ROLE		**699**	**443**	**82**	**84.4**	**63.4**	**72.4**
	Citizen-Of	36	25	8	75.8	69.4	72.6
	General-Staff	201	108	46	71.1	53.7	62.3
	Management	165	106	72	59.6	64.2	61.8
	Member	224	104	36	74.3	46.4	57.1
SOCIAL		**95**	**60**	**21**	**74.1**	**63.2**	**68.5**
	Other-Professional	29	16	32	33.3	55.2	41.6
	Parent	25	17	0	100	68.0	81.0

Table 4: Performance of different relation types and major subtypes in the test data

Supervised RE: summary

- Supervised approach can achieve high accuracy
  - At least, for *some* relations
  - If we have lots of hand-labeled training data
- But has significant limitations!
  - Labeling 5,000 relations (+ named entities) is expensive
  - Doesn't generalize to different relations
- Next: beyond supervised relation extraction
  - Distantly supervised relation extraction
  - Unsupervised relation extraction